

# 应用于材料图像分割的 Graph-UNet<sup>\*</sup>

魏惠珊<sup>1</sup>, 韩越兴<sup>1,2†</sup>, 王 冰<sup>1</sup>, 陈侨川<sup>1</sup>

(1. 上海大学 计算机科学与工程学院, 上海 200444; 2. 之江实验室, 杭州 311100)

**摘 要:** 小样本材料图像分割是图像分割领域的研究难点之一。材料图像的微观结构大多数有形状各异、纹理复杂和边界模糊等特点, 会导致材料图像的分割不准确。Graph-UNet 被提出融合 UNet 和图卷积神经网络来解决小样本材料图像自动分割的挑战, 它将卷积神经网络的多维特征融合和跳跃连接的思想迁移到图卷积神经网络中实现图卷积和图注意力的有效结合, 并且建立了一个通用的模块实现特征图和图结构相互转换。在材料图像数据集上进行了对比和消融实验, 证明 Graph-UNet 的分割结果优于很多先进方法, 准确的识别了多种材料结构, 推动了探究材料结构和性能关系的发展。

**关键词:** 语义分割; 图卷积神经网络; 图注意力; 材料图像; 深度学习; 小样本

**中图分类号:** TP183      **doi:** 10.19734/j.issn.1001-3695.2022.03.0132

## Graph-UNet for material image segmentation

Wei Huishan<sup>1</sup>, Han Yuexing<sup>1,2†</sup>, Wang Bing<sup>1</sup>, Chen Qiaochuan<sup>1</sup>

(1. School of Computer Engineering & Science, Shanghai University, Shanghai 200444, China; 2. Zhejiang Laboratory, Hangzhou 311100, China)

**Abstract:** Image segmentation of small sample material is one of the difficulties in the field of image segmentation. Most of the microstructure of material image has the characteristics of different shape, complex texture and fuzzy boundary, which will lead to the inaccurate segmentation of material image. Graph-UNet is proposed to integrate UNet and Graph convolutional neural network to solve the challenge of automatic image segmentation of small sample materials. The model transferred the idea of multi-dimensional feature fusion and jump connection from convolutional neural network to graph convolutional neural network. The model realizes the effective combination of graph convolution and graph attention. A universal module is established to convert feature graph and graph structure to each other. The comparison and ablation experiments on material image data sets prove that Graph-UNet segmentation results are superior to many advanced methods and can accurately identify multiple material structures, which promotes the development of exploring the relationship between material structure and properties.

**Key words:** semantic segmentation; graph convolutional neural network; graph attention; material image; deep learning; small sample

## 0 引言

随着图像分割方法的不断发展, 以及各领域间的交叉和融合, 科学工作者尝试用先进的图像分割方法运用到材料图像分割中, 如文献[1,2]。材料图像的精准分割有助于推动材料的制造工艺、结构和性能之间的关系的研究, 有助于新材料的开发。材料结构的图像分割是对材料图像进行后续处理的基础。材料图像中目标区域的准确分割对于结构分析、研究材料性能等都具有重要意义。

大多数材料图像包含复杂的纹理和各种形状的微观结构, 即使是同种材料结构也会存在差别很大的形状和纹理。材料结构千差万别, 材料图像的标注需要耗费大量的人力和时间成本。因此提高模型的泛化性和解决小样本图像分割问题也是必不可少的。在小样本数据集中, 区分不同的材料结构图像是一个困难的任务。

目前已经有了与材料图像分割有关的深度学习方法。Azimi 等人<sup>[3]</sup>利用全卷积神经网络(FCN)提出了一种新型的针对材料显微组织的分割算法 MVFCNN。该方法需要较多的

材料样本, 由于大多数材料图像的获取成本过高, 这限制了深度学习在材料图像分割中的应用。Jiang 等人<sup>[4]</sup>提出了一种基于深度学习的方法, 直接对整幅图像进行像素级材料分割, 将扩张卷积特征与传统卷积特征相结合, 去除扩张卷积带来的伪影。Xiong 等人<sup>[5]</sup>提出了一种基于简单线性迭代聚类超像素区域生成和卷积神经网络分类的水稻穗段分割算法。然而, 水稻穗段有着明显的边界和区域特征, 使用超像素方法存在特定任务的局限性。Decost 等人<sup>[6]</sup>提出一种深度卷积神经网络, 能够从含有多种微观成分的微观图中获得水石灰粒度和变质区宽度的分布。然而, 他们主要是针对特定的应用场景设计网络模型。

为了解决端到端的材料图像分割任务, 本文结合 UNet 和图卷积神经网络提出了一个通用的网络模型, Graph-UNet, 如图 1 所示。本文利用深度卷积网络对一个图进行建模, 并且应用图卷积方法来解决材料图像的语义分割任务。该模型将卷积神经网络中处理的特征图建立一个图结构, 使用多维度跳跃连接的图注意力模型(MGAM)在邻居节点之间交换特征, 再将交换信息后的图结构封装为特征图进行后续的卷积

**收稿日期:** 2022-03-15; **修回日期:** 2022-05-17      **基金项目:** 国家重点研发计划项目(2018YFB0704400, 2020YFB0704500); 上海市自然科学基金项目(20ZR1419000); 之江实验室科研攻关项目(2021PE0AC02)

**作者简介:** 魏惠珊(1997-), 女, 安徽合肥人, 硕士研究生, 主要研究方向为材料图像分割; 韩越兴(1975-), 男(通信作者), 黑龙江海伦人, 副教授, 博导, 博士, 主要研究方向为人工智能、图像处理(han\_yx@i.shu.edu.cn); 王冰(1978-), 女, 黑龙江牡丹江人, 教授, 博导, 博士, 主要研究方向为复杂网络理论方法、数据处理; 陈侨川(1991-), 男, 安徽合肥人, 讲师, 硕导, 博士, 主要研究方向为图像处理。

层操作。本文方法主要有两点贡献: a)提出了一种新型的图注意力模块, 将图卷积和图注意力进行多种不同维度的特征融合; b)提出了一个应用于材料图像的图模型, 将图像封装成图结构参与图卷积。

本文的其余部分内容如下: 在第一节介绍了本文的网络框架和具体细节。第二节展示了在材料数据集上的实验结果和讨论。第三节是对本文工作的总结和对未来的展望。

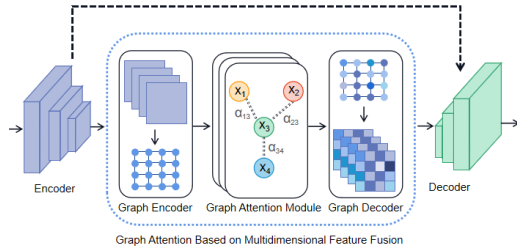


图 1 网络结构

Fig. 1 Network structure

## 1 语义分割中的图模型

图卷积神经网络是为解决在图结构数据集上的学习问题而设计的。如今, 研究者开始探索图卷积如何应用在结构化数据上。Lu 等人<sup>[7]</sup>首次用图卷积结合全卷积神经网络解决语义分割难题, 使模型有 1.34% 的性能提升。Zhang 等人<sup>[8]</sup>提出对偶图卷积网络(DGCNet), 将处理不规则数据的图卷积融合到解决语义分割任务的双重注意力模型中。Ma 等人<sup>[9]</sup>提出了一种新的注意图卷积网络(AGCN)来对遥感图像数据进行超像素分割, 利用图卷积结合图注意力实现图像分割。

目前, 很少有结合图卷积的网络模型针对材料图像进行分割, 借鉴前人的研究思路, 本文设计图结构主要有两个特点。首先, 图结构有更加灵活的跳跃连接, 不像卷积网络中以一个像素为中心, 建立一个  $n \times n$  的局部区域进行卷积, 图卷积中, 每个节点的领域可以自定义范围。其次, 图卷积的消息传播机制在节点间交换信息, 随着图卷积层的堆叠, 每个节点的感受野变得更大, 充分捕捉更大范围的特征。

为了延续注意力机制在图像分割任务中的优点, 增强网络对重要特征的专注, 本文将基于多维特征融合的图注意力模块加入 UNet<sup>[10]</sup>中解决小样本材料图像的语义分割问题。本文选择 UNet 作为骨干网络的初衷是因为 UNet 最初应用于医学图像, 而材料图像与医学图像有很大的相似之处, 例如颜色单调, 以灰度图为主。现在, 对于 UNet 的改进方法经典的有 UNet++<sup>[11]</sup>和 Attention-UNet<sup>[12]</sup>等。UNet++主要通过增加跳跃连接增强对图像特征的补充, 而 Attention-UNet 通过在解码器部分增加注意力机制提高网络性能。本文受到图卷积和注意力机制在图像分割任务中应用的启发, 提出了图注意力模块。

在本文提出的网络中, 输入图像首先经过五层卷积层, 前四层随后各加一层池化层。在第五层卷积层之后, 本文将提取的特征图转换为图结构, 用基于多维特征融合的图注意力模块进行节点特征的传播和学习。随后, 模块将图结构再转换回特征图进入后面的四层上采样层, 不断地将特征图的尺寸放大。每一次上采样操作后都会将当前的特征图和编码器相同尺寸大小的特征图进行拼接。最后, 网络通过 Sigmoid 函数输出预测图, 得到分割结果。

### 1.1 图结构

图结构是一种非规则的数据结构。图结构数据可以视为一个三元组  $G(N, E, U)$ 。N 是图的节点集, 它是一个  $|N| \times S$  矩阵,  $|N|$  是图的节点数,  $S$  是节点特征向量的维数,  $E$  是图的边集,  $U$  是图的特征。

本文采用有向图的方式构建图, 即对每条边赋予方向性。为了实现图结构的建立后, 每个节点之间的连接情况没有差别性, 本文会将每个节点的总度初始化为特定值, 即四邻域的方式。本文通过监督学习来训练这个模型。

### 1.2 基于多维特征融合的图注意力

#### 1.2.1 图编码器

图注意力模块的输入是图结构。为了将特征图输入到图注意力模块, 本文创建了图编码器把特征图转换为图结构。

在本文模型中, 节点特征由 UNet 的编码器部分初始化。在图模型中, 用邻接矩阵表示节点间边的连接, 每个节点都连接到它最近的节点。本文根据像素点的空间位置关系让每个节点按照四邻域的方式建立边的连接。节点间的连接表示特征可以通过图注意力模块中的边进行传播。

#### 1.2.2 图注意力模块

本文提出了一种基于多维特征融合的图注意力模块(MGAM), 如图 2 所示。

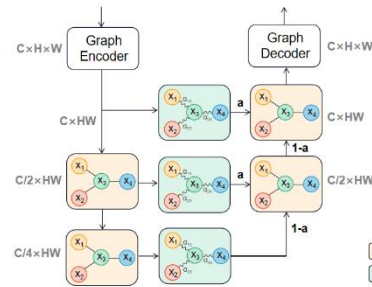


图 2 图注意力模块

Fig. 2 Graph attention module

图卷积神经网络(GCN)是处理图结构的深度学习方法之一, 它的具体实现公式<sup>[13]</sup>如下:

$$Z^{l+1} = \sigma(D^{-\frac{1}{2}} A D^{-\frac{1}{2}} X W^l) \quad (1)$$

$X$  是由 UNet 编码器的输出转换的图结构。 $D$  和  $A$  分别是  $X$  的度矩阵和邻接矩阵。 $\sim$  是指图的每个节点添加了自连接, 并构建了相应的度矩阵  $D$  和邻接矩阵  $A$ 。 $W^l$  是第  $l$  层图卷积的权重矩阵。 $\sigma$  是 ReLU 激活函数。 $Z^{l+1}$  是第  $l+1$  层图卷积操作的输出。

图注意力网络(GAT)是应用于图结构的注意力机制之一, 它的具体实现公式<sup>[14]</sup>如下:

$$\bar{h}_i(K) = \sigma\left(\frac{1}{K} \sum_{k=1}^K \sum_{j \in N_i} \alpha_{ij}^k W_k^T \bar{h}_j\right) N_i \quad (2)$$

$K$  表示注意力头的个数,  $N_i$  是节点  $i$  的邻居节点集合,  $\alpha_{ij}^k$  是第  $k$  个注意力头中节点  $i$  和它的第  $j$  个邻居节点之间的注意力系数,  $W_k$  是第  $k$  个注意力头的权重矩阵,  $h_j$  是第  $j$  个节点的特征向量,  $\sigma$  是激活函数,  $\bar{h}_i$  是节点  $i$  经过多头图注意层的聚合后的输出特征。其中  $\alpha_{ij}$  的具体公式如下:

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(a^T [W_h \| W_l]))}{\sum_{r \in N_i} \exp(\text{LeakyReLU}(a^T [W_h \| W_l]))} \quad (3)$$

LeakyReLU() 是特定的激活函数,  $a$  和  $W$  都是权重矩阵,  $h_i$ 、 $h_j$  和  $h_r$  分别是节点  $i$ 、节点  $j$  和节点  $i$  的第  $r$  个邻居节点。

GCN 和 GAT 都是将邻居节点的特征聚合到中心节点, GCN 利用了拉普拉斯矩阵, GAT 利用了注意力系数。GCN 不能为节点的每个邻居分配不同的权重, 在卷积时对所有邻居节点均一视同仁, 不能根据节点重要性分配不同的权重。然而, GAT 是将节点特征之间的相关性很好的融合到了模型。GAT 利用多头注意力以稳定学习过程, 如图 3 所示。

一般来说, GCN 解决节点分类或分割问题采用一到两层, 过多层的 GCN 堆叠会导致过平滑的问题。然而, 卷积神经网络



络模型中,一般深层的卷积层堆叠会展现更好的性能和效果。更深的模型,意味着更好的非线性表达能力,可以学习更加复杂的变换,从而可以拟合更加复杂的特征输入,例如残差网络。为了借鉴 CNN 的思想,本文通过增加超参数  $a$  控制每层节点特征输入下一层图卷积层的比例,加深了图卷积模型的深度。这种思想模拟了残差网络的跳跃连接,但这种连接仅部分缓解了过度平滑的问题。当堆叠更多层时,模型的性能仍然会下降。

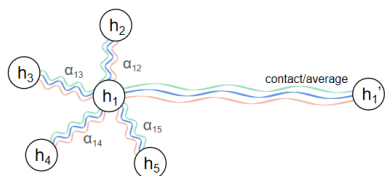


图 3 图注意力的多头注意力机制

Fig. 3 The multi-attentional mechanism of graph attention

本文使用两层 GCN 一边对节点特征进行消息传播一边对节点特征降维操作,再通过两层 GCN 对节点特征升维操作且加和每次降维的节点特征。首先,输入图结构有  $H \times W$  个节点,每个节点特征维度为  $C$ 。经过一层图卷积层,将节点特征的维度降到  $C/2$ 。然后,再将节点特征更新过一次的图输入第二层图卷积层,节点特征维度降为  $C/4$ 。紧接着,对图进行升维操作,分别再进入两层图卷积层。降维和升维后的特征一一对应,并且降维后的节点特征连接一层图注意力 GAT 再加和到升维的图卷积层。维度为  $C$ 、 $C/2$  和  $C/4$  的图结构分别对应图像领域中的高分辨率、中分辨率和低分辨率的特征图。通过代表低分辨率特征图的图结构学习到粗糙的特征,代表高分辨率特征图的图结构学习到精细的细节特征,将不同分辨率的图结构融合达到对小样本材料图像的准确学习和表示。在图经过图注意力层后,通过设置一个超参数  $a$  控制节点特征流动到下一层图卷积层的比例。设置超参数  $a$  的目的是为了避免过多的图卷积层导致节点特征最终都几乎倾向等于全局特征。特征升维部分的具体融合实现方式如式(4)所示。

$$H_{add}^{l+1} = a\bar{h} + (1-a)Z^{l+1} \quad (4)$$

$H_{add}^{l+1}$  是图注意力层输出和图卷积层升维操作的输出的加和结果,  $Z^{l+1}$  是第  $l+1$  层图卷积的输出特征,  $a$  是超参数,  $\bar{h}$  是图注意力层的输出。

图注意力模块是网络模型的重要组成模块,是图卷积和图注意力组合的结果。它主要起到了加深网络深度,提高网络拟合复杂的输入图像的能力,缓解网络对局部特征的过度关注,避免了忽视全局特征信息,达到更好的材料图像分割效果。图卷积不会改变特征图像素点的空间位置信息,节点不会消失。图注意力层起到了关注重要节点的作用,增强重要节点的特征流动,抑制无关节点的特征流动。

### 1.2.3 图解码器

图注意力模块的输出依然是图结构。为了将图注意力模块的输出可以继续放入卷积神经网络中,本文创建了解码器将图结构转换为特征图。

## 2 实验和讨论

### 2.1 实验环境与评价标准

在实验中,本文使用了 i7-8750H CPU@2.20Hz 的计算机,以及 Python3.8 的开发工具。网络是通过 PyTorch 实现的。

本文将图片大小统一为  $512 \times 512$  的大小作为网络输入,并且将标签图像转换为二值化图像用于训练阶段的使用,将二元交叉熵损失函数作为网络的损失函数。本文使用带动量

的梯度下降优化算法 Adam 训练网络模型,基于训练数据迭代地更新神经网络权重;动量和衰减系数为 0.9 和 0.0005,以减轻模型过拟合的问题;迭代次数为 100,批处理大小为 1,初始学习率设为 0.001,通过设置学习率逼近网络参数最优值。每种材料结构的数据集包含两类标签,即目标材料结构和背景材料结构。本文将训练数据集放入网络训练,通过损失函数降低输入图像和标签图像之间的损失,保存 100 次迭代中网络性能最优的网络参数以备测试阶段的使用。网络性能是否优良的评价标准是使用 MIoU 来衡量。

本文将训练数据集放入网络训练,保存训练最优的网络参数,再放入测试数据集进行预测,计算测试集预测结果与对应标注图的 MIoU 作为评价网络性能的指标。最后,本文将网络的输出结果转换为可视化的二值化标注结果图。

本文实验的数据集一共四类材料结构,分别是 Spheroidite、Wood、Pearlite1 和 Pearlite2。其中, Spheroidite、Pearlite1 和 Pearlite2 是开源的材料数据,来源于文献[15]。材料数据集由于实验复杂和标注成本高昂,每类有 4 到 6 张图片,数据集分布如表 1 所示。

表 1 数据集分布

Tab. 1 Data set distribution

class	spheroidite	wood	pearlite1	pearlite2
train set	4	4	4	3
test set	2	2	2	1
data set	6	6	6	4

本文主要用 MIoU 作为语义分割网络的评价指标,具体计算公式如下:

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (5)$$

$p_{ii}$  表示真实值为  $i$ , 被预测为  $i$  的数量。 $p_{ij}$  表示真实值为  $i$ , 被预测为  $j$  的数量。 $p_{ji}$  表示真实值为  $j$ , 被预测为  $i$  的数量。 $k$  是类别个数。MIoU 一般都是基于类进行计算的,将每一类的 IoU 计算之后累加,再进行平均,得到的就是基于全局的评价。

### 2.2 与现有方法的对比实验

为进一步证明提出的模型能有效的应用在材料图像分割任务中,本文比较了在同样数据集中当前先进的方法。

本文比较了传统的图像分割方法,包括 KMeans<sup>[16]</sup>和 Watershed<sup>[17]</sup>的分割结果图,还有深度学习方法 FCN<sup>[18]</sup>、

UNet<sup>[10]</sup>、UNet++<sup>[11]</sup>、CENet<sup>[19]</sup>、R2U-Net<sup>[20]</sup>、Attention-UNet<sup>[12]</sup>和 Unet3++<sup>[21]</sup>。实验结果如图 4 和表 2 所示, KMeans 对复杂纹理图像中的噪声过于敏感,导致绝大多数纹理的细节特征都被分割出来。因此, KMeans 不能得到完整的材料结构,各相被过度分割。Watershed 的分割结果层次不齐,如图 4 中 Watershed 第 5、6 和 7 图,基本没有分割出目标结构甚至出现了分割图大面积呈现黑色的错误结果。可见对于对比度不强烈的灰度图, Watershed 难以实现较好的分割效果。KMeans 和 Watershed 在图像分割任务上缺乏通用性。各个基于深度学习的网络模型在材料图像分割任务上的表现因为样本匮乏,无法得到丰富的数据集训练,导致对材料结构的特征学习十分不充足。FCN 的实验结果出现了不同程度的噪声。见图 4 中 FCN 的第 3 张图,边界附近出现了许多微小但明显的噪声,说明 FCN 对于边界附近的特征学习还有所欠缺。见图 4 中 FCN 的第 7 张图,在块状的材料结构存在明显的颜色差异,即中心趋向黑色,边缘呈现白色。对于这样的材料结构, FCN 没有很好的将属于同一块的结构分割为一块,分割出很多不存在的边界。显然, FCN 没有很好的结合上下文信息,导致分割出很多小区域。UNet、UNet++和 UNet+++

的分割结果都不错,但是三个网络相比之下,UNet++的分割可视化效果略欠一筹,出现了一些噪声,这说明对于材料数据集来说过多的跳跃连接不一定是百利而无一害。CENet 的评价指标和分割图都趋于稳定,但依然存在小范围的分割错误,如图 4 中 CENet 的第 2 张和第 7 张图。R2U-Net 的实验结果出现了局部分割错误。Attention-UNet 的分割结果也出现了层次不齐。更详细的统计如表 2 所示,本文的方法在 MIoU 指标上也表现的更加优异,大多超过现有方法 2~5 个百分点左右。

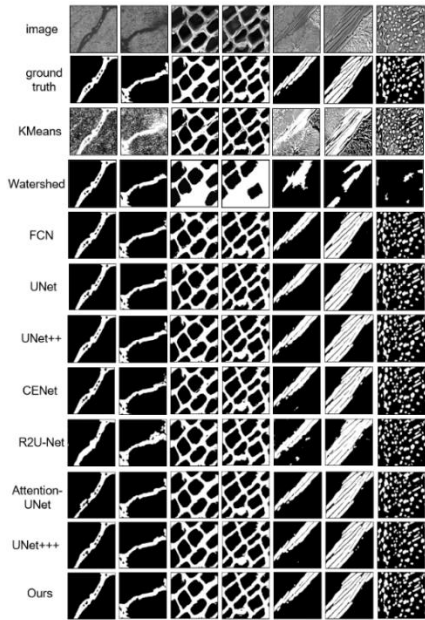


图 4 对比实验

Fig. 4 Contrast experiment

表 2 对比实验

Tab. 2 Comparative variety trial

Methods	spheroidite	wood	pearlite1	pearlite2	aver.
W <sup>[12]</sup>	41.5487	88.4293	32.3180	58.7838	55.2700
	91.8501	63.6468	58.7721	35.0380	62.3268
	94.3108	88.3342	92.4786	88.1496	90.8183
	95.6969	89.3567	93.7345	91.6389	92.6068
	96.1455	89.9543	94.1003	91.4282	92.9071
	94.0800	89.0585	93.3075	91.3135	91.9399
	83.2993	88.2415	85.8098	87.0114	86.0905
	91.9393	91.2300	94.8990	89.8639	91.9831
	92.8708	89.2140	93.6977	87.5113	90.8235
	96.9200	90.3073	94.6179	91.5946	93.3600
本文 approach	96.9200	90.3073	94.6179	91.5946	93.3600

本文的网络模型对多类数据集的分割效果呈现稳定的性能。材料结构的大部分噪点都被很好的过滤,对大面积的连通区域分割准确。材料结构之间的狭窄边界分割优于绝大多数方法。在颜色对比度低的情况下,依然展现强大的学习能力。本文的方法较好的权衡了分类准确率和定位精度。

2.3 图注意力模块有效性的验证实验

本文探究了基于不同 backbone 的网络的性能,即验证图注意力模块的有效性,对比各个网络模型添加图注意力模块和不添加图注意力模块的网络性能。本文对比了 backbone 为 ResNet34-UNet、ResNet34-FCN 和 UNet++ 的网络模型,发现图注意力模块对于提高网络性能有着积极的作用,如表 3 所示。

为了证明提出的图注意力模块对于材料图像分割的积极作用,对比了以往经典的注意力方法,即 SE-Net<sup>[22]</sup>、CBMA<sup>[23]</sup>、DANet<sup>[24]</sup>、CCNet<sup>[25]</sup>和 PSANet<sup>[26]</sup>,如表 4 所示。本文把以上五种注意力方法作为五个不同的模块和提出的图注意力模块分别添加到 UNet 中在四个数据集上进行验证对比。实验证明提出的模块对于材料图像的分割有着积极影响。

表 3 不同骨干网络的对比实验

Tab. 3 Comparative experiments of different backbone networks

Methods	spheroidite	wood	pearlite1	pearlite2	Mean
R <sub>34</sub> -UNet-MGAM	95.0883	88.0642	92.6186	90.5320	91.5758
	95.4993	88.5567	93.6210	91.0782	92.1888
	94.1122	88.8613	93.4565	90.9899	91.8550
R <sub>34</sub> -FCN-MGAM	95.5784	89.5392	94.0044	91.1701	92.5730
	96.1455	89.9543	94.1003	91.4282	92.9071
	96.5721	90.0362	94.1026	91.4533	93.0411
UNet++-MGAM	95.6969	89.3567	93.7345	91.6389	92.6068
	96.9200	90.3073	94.6179	91.5946	93.3600
本文 approach	96.9200	90.3073	94.6179	91.5946	93.3600

表 4 不同注意力方法的对比实验

Tab. 4 A comparative experiment of different attention methods

Methods	spheroidite	wood	pearlite1	pearlite2	Mean
UNet	93.4806	88.3570	93.1956	89.8086	91.2105
	96.3472	88.2062	93.3316	90.3746	92.0649
	96.3901	88.3589	94.5993	90.9791	92.5819
	95.8607	88.4541	92.0037	88.0546	91.0933
	94.6511	89.4653	93.9191	90.2729	92.0771
本文 approach	96.9200	90.3073	94.6179	91.5946	93.3600

2.4 消融实验

在本文的工作中,主要提出了一个新颖的基于多维特征融合的图注意力模块。为了了解图卷积、图注意力和超参数  $a$  对网络模型的影响,本文做了详细的消融实验。本文主要做了五组实验。第一组的网络模型是 UNet 加一层图卷积层。第二组是 UNet 加一层图注意力层。第三组是 UNet 加上基于多维特征融合的图卷积模块。第四组是 UNet 加上没有超参数  $a$  的基于多维特征融合的图注意力模块。第五组是本文的网络模型。对比了五组不同的网络模型在材料数据集上的分割结果,不同图卷积模块的结构如图 5 所示。

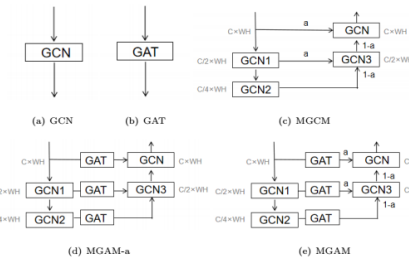


图 5 不同图卷积模块的结构

Fig. 5 Structure of different graph convolution modules

消融实验的结果如表 5 所示,报告中的数字以百分比作为分割精度。从表格中发现,本文提出的基于多维特征融合的图注意力模块对于模型性能有着积极的影响,图注意力模块中的各个部分对于分割效果都有着不同程度的提升作用。实验的可视化结果如图 6 所示,每一组的网络模型对于材料结构的分割结果都是较为精确的,本文难用肉眼去发现不同组之间的区别。然而,在边界的分割中,发现在每一组实验的第五张和第六张图,本文提出的模块对于识别边界有很好的积极影响。由此可得,基于多维特征融合的图注意力模块提升了对于微小细节的分割效果。

表 5 消融实验

Tab. 5 Ablation experiments

Methods	spheroidite	wood	pearlite1	pearlite2
UNet+GCN	96.2792	89.1921	94.2530	90.0736
UNet+GAT	92.6897	88.3697	93.9177	92.3606
UNet+MGCM	95.8542	90.1897	94.5596	90.2766
UNet+MGAM-a	96.4528	89.8272	93.8442	89.9251
本文 approach	96.9200	90.3073	94.6179	91.5946

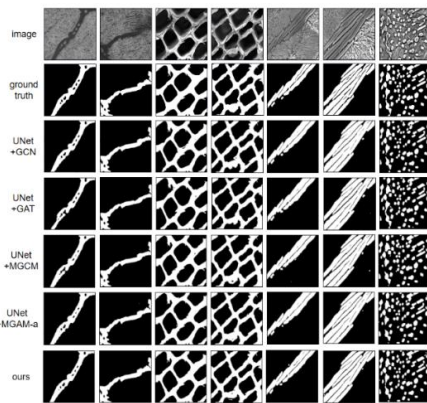


图 6 消融实验

Fig. 6 Ablation experiments

### 3 结束语

在这项工作中, 提取了一种针对小样本材料结构图像的语义分割方法。本文将基于多维特征融合的图注意力模块放入 UNet 中, 在样本数量少的前提下获得更多的特征信息去实现对材料图像的自动分割。该方法有一些缺点有待改进。例如本文方法在计算成本方面相对昂贵。其次, 材料结构不明显的边界细节还是不能很好的分割出来。对于未来的研究工作, 将进一步加强本文方法, 如结合图卷积的优点使用更深的图卷积神经网络提高模型的拟合能力, 以应用于更多的新材料结构。

### 参考文献:

- [1] 何栋, 唐婷. 基于模糊 C 均值聚类算法的碳纤维木质复合材料图像分割 [J]. 合成材料老化与应用, 2020 (4): 78-80. (He Dong, Tang Ting. Image segmentation of carbon fiber wood composite based on fuzzy C-means clustering algorithm [J]. Synthetic Materials Aging and Application, 2020 (4): 78-80.)
- [2] 张利欣, 车世界, 徐正光, 等. 基于残差网络的高温合金微观组织图像分割方法 [J]. 科学技术与工程, 2020, 20 (1): 246-251. (Zhang Lixin, Che Shijie, Xu Zhengguang, et al. Image segmentation method of superalloy microstructure based on residual network [J]. Science Technology and Engineering, 2020, 20 (1): 246-251.)
- [3] Azimi S, Mücklich F, Fritz M. Advanced steel microstructural classification by deep learning methods [J]. ArXiv: 1706. 06480, 2018.
- [4] Jiang X, Du J, Sun B, et al. Deep dilated convolutional network for material recognition [C]. 2018 Eighth International Conference on Image Processing Theory, Tools and Applications (IPTA), 2018. [http://dx. doi. org/10. 1109/IPTA. 2018. 8608160](http://dx.doi.org/10.1109/IPTA.2018.8608160).
- [5] Xiong X, Duan L, Liu L, et al. Panicle-SEG: A robust image segmentation method for rice panicles in the field based on deep learning and superpixel optimization [J]. Plant Methods, 2017, 13 (1): 104-118.
- [6] Decost B L, Lei B, Francis T, et al. High throughput quantitative metallography for complex microstructures using deep learning: A case study in ultrahigh carbon steel [J]. ArXiv: 1805. 08693, 2018.
- [7] Lu Y, Chen Y, Zhao D, et al. Graph-FCN for image semantic segmentation [J]. Advances in Neural Networks, 2019. [https://doi. org/10.](https://doi.org/10.1007/978-3-030-22796-8_11)

- 1007/978-3-030-22796-8\_11.
- [8] Huang Z, Wang X, Huang L, et al. CCNet: Criss-cross attention for semantic segmentation [C]. International Conference on Computer Vision, 2019. [http://dx. doi. org/10. 1109/ICCV. 2019. 00069](http://dx.doi.org/10.1109/ICCV.2019.00069).
- [9] Ma F, Gao F, Sun J, et al. Attention graph convolution network for image segmentation in big SAR imagery data [J]. Remote Sensing, 2019, 11 (21): 2586-2606.
- [10] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation [J]. ArXiv: 1505. 04597, 2015.
- [11] Zhou Z, Siddiquee M, Tajbakhsh N, et al. UNet+: A nested U-Net architecture for medical image segmentation [J]. ArXiv: 1807. 10165, 2018.
- [12] Oktay O, Schlemper J, Folgoc L L, et al. Attention U-Net: Learning where to look for the pancreas [J]. ArXiv: 1804. 03999, 2018.
- [13] Kip F, M. Welling. Semi-supervised classification with graph convolutional networks. ArXiv: 1710. 10903, 2016.
- [14] Velickovi P, Cucurull G, Casanova A, et al. Graph attention networks [J]. 2017.
- [15] Decost B L, Hecht M, Francis T, et al. UHCSDB: UltraHigh carbon steel micrograph database [J]. Integrating Materials&Manufacturing Innovation, 2017, 6 (8): 1-9.
- [16] Macqueen J. Some methods for classification and analysis of multiVariate observations [C]. In 5-th Berkeley Symposium on Mathematical Statistics and Probability, 1965, 66 (1): 281-297.
- [17] Bieniek A, Moga A. An efficient watershed algorithm based on connected components [J]. Pattern Recognition, 2000, 33 (6): 907-916.
- [18] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39 (4): 640-651.
- [19] Gu Z, Cheng J, Fu H, et al. CE-Net: Context encoder network for 2D medical image segmentation [J]. IEEE Transactions on Medical Imaging, 2019, 38 (10): 2281-2292.
- [20] Alom M Z, Hasan M, Yakopcic C, et al. Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation [J]. ArXiv: 1802. 06955, 2018.
- [21] Huang H, Lin L, Tong R, et al. UNet 3+: A full-scale connected UNet for medical image segmentation [J]. ArXiv: 2004. 08790, 2020.
- [22] Jie H, Li S, Gang S, et al. Squeeze-and-excitation networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 42 (8): 2011-2023.
- [23] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module [J]. ArXiv: 1807. 06521, 2018.
- [24] Fu J, Liu J, Tian H, et al. Dual attention network for scene segmentation [C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020. [http://dx. doi. org/10. 1109/CVPR. 2019. 00326](http://dx.doi.org/10.1109/CVPR.2019.00326).
- [25] Huang Z, Wang X, Huang L, et al. CCNet: Criss-cross attention for semantic segmentation [C]. International Conference on Computer Vision, 2019. [http://dx. doi. org/10. 1109/ICCV. 2019. 00069](http://dx.doi.org/10.1109/ICCV.2019.00069).
- [26] Liu H, Liu F, Fan X, et al. Polarized self-attention: towards high-quality pixel-wise regression [J]. ArXiv: 2107. 00782, 2021.